LA-UR-18-29734
Approved for public release; distribution is unlimited.

Title:	Advances in machine learned potentials for molecular dynamics simulation
Author(s):	Barros, Kipton Marcos
Intended for:	Release slides online
Issued:	2018-10-13

Advances in machine learned potentials for molecular dynamics simulation

Kipton Barros
Los Alamos National Lab.

Physics Next, Machine Learning,
Oct 9, 2018, Riverhead, NY

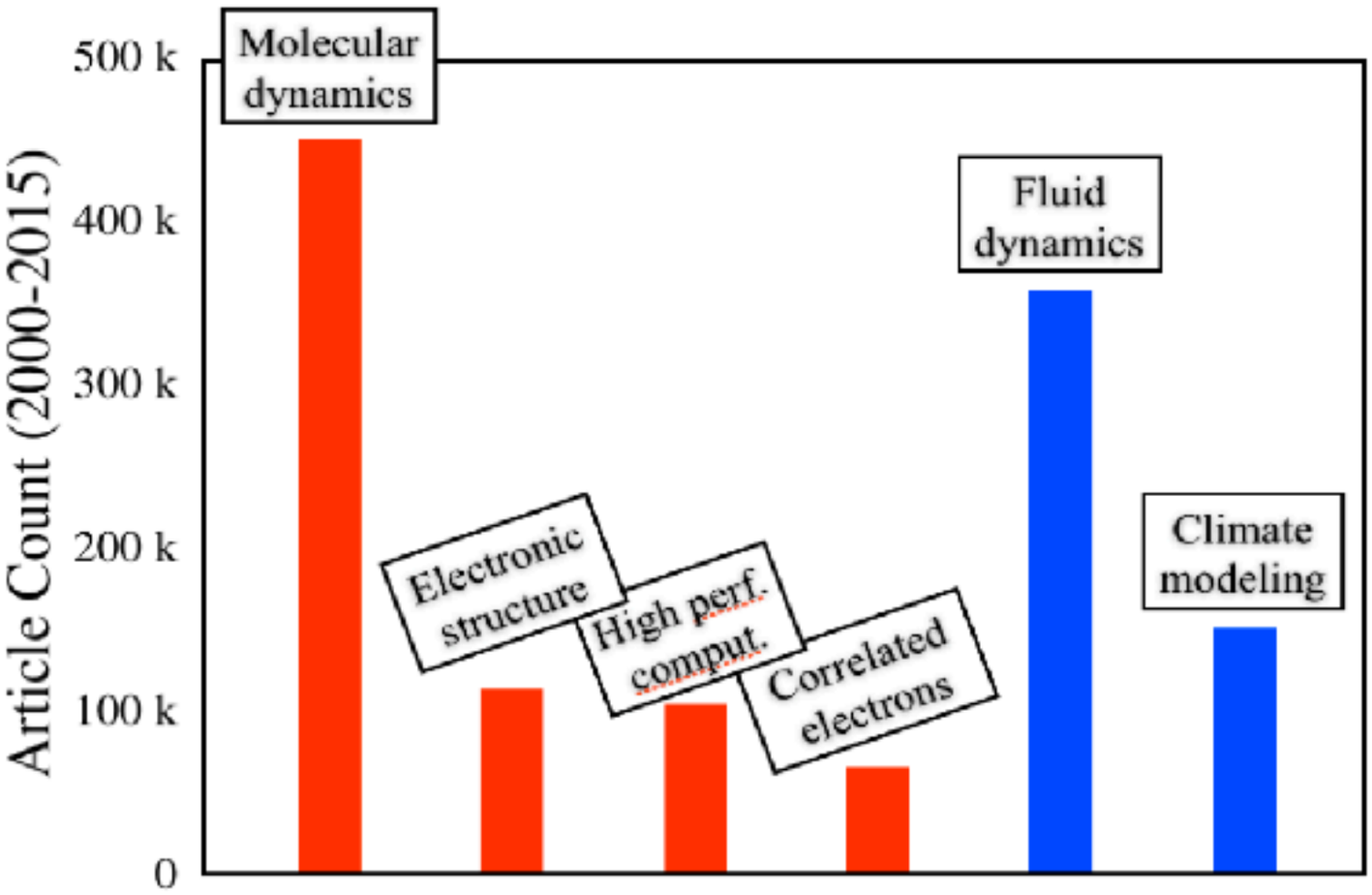
Collaborators

LANL:

Nick Lubbers
Justin Smith
Andrew Sifain
Ben Nebgen
Sergei Tretiak

Oles Isayev (UNC)
Adrian Roitberg (UF)

Molecular dynamics articles

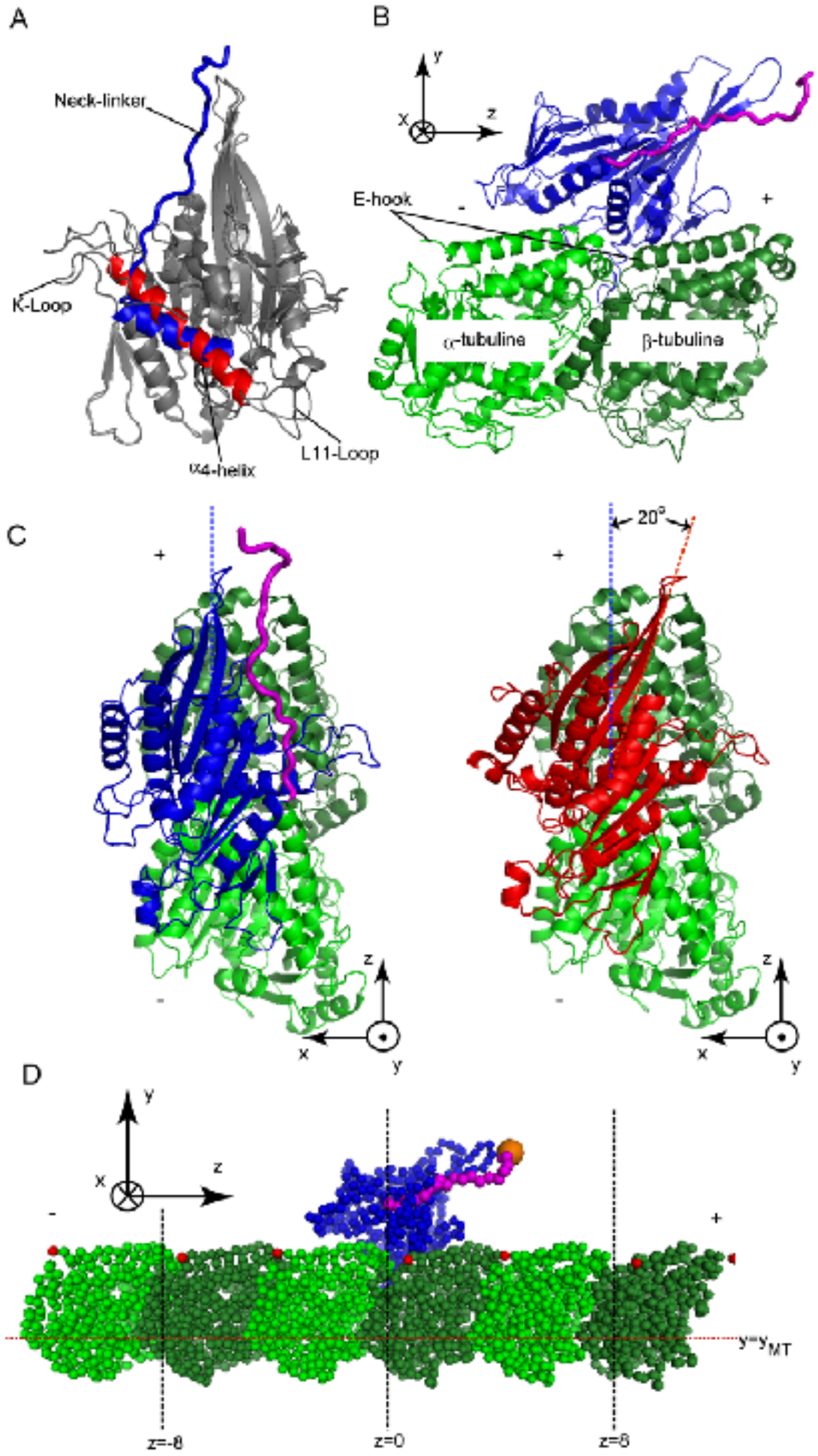


Drug design

Biophysics

Materials

...



https://doi.org/10.1371/journal.pcbi.1002907

Born Oppenheimer MD

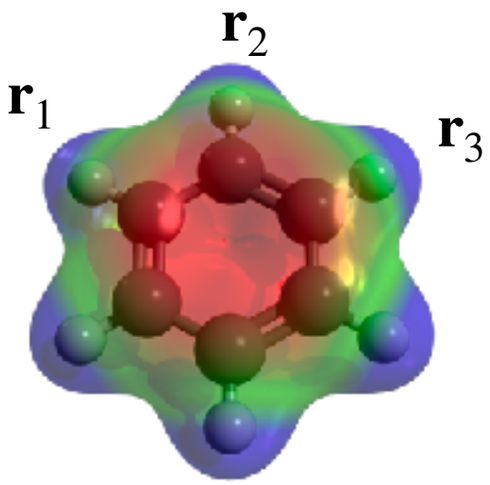
→ separation of scales

Calculate energy/forces for fixed nuclei positions

$$\mathbf{r}_1 \ldots \mathbf{r}_N \mapsto E[\mathbf{r}]$$

Schrödinger Eq. ≈

{Hartree Fock, DFT,
Coupled Cluster, …}



Integrate nuclei dynamics

$$\frac{d^2\mathbf{r}_i}{dt^2} = -\nabla_i E[\mathbf{r}]/m_i$$

Not treated:

Quantum mechanics of nuclei (e.g. hydrogen)

Non-equilibrium dynamics of electrons

- von Neumann eq. for density matrix
- non-adiabatic excited state MD

Classical potential / force field

Pros: Fast

Cons: Not very transferable, Non-reactive
Laborious parameterization

Schrödinger Equation

Pros: Accurate, transferable

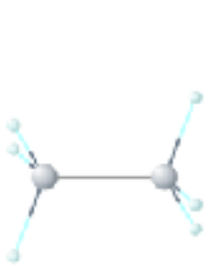
Cons: Computationally demanding

Machine learning to emulate Schrödinger eq.

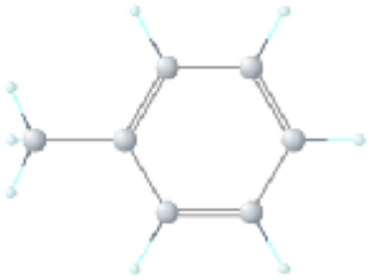
QM9 - 130k organic molecules

Relaxed geometries

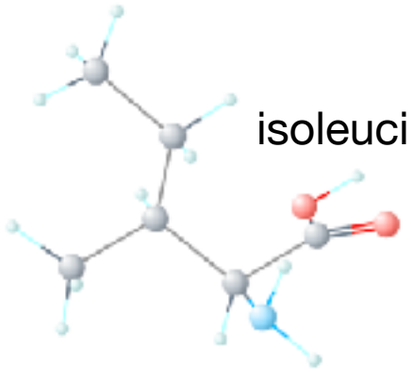
C, N, O, F (up to 9) + H



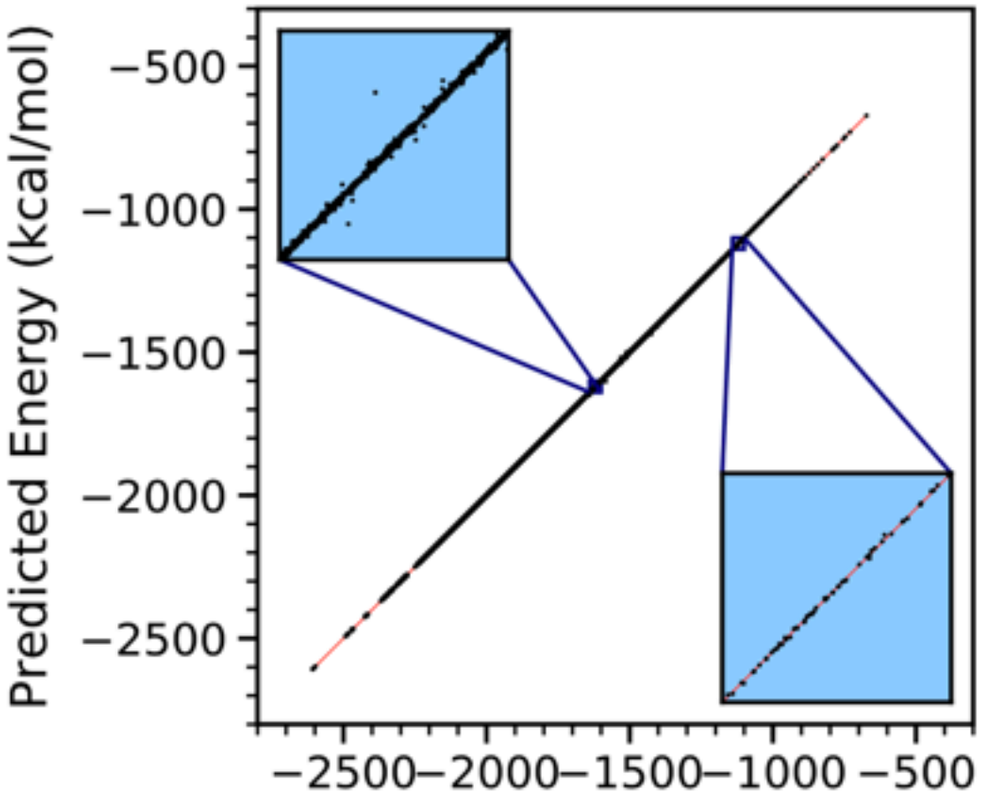
ethane



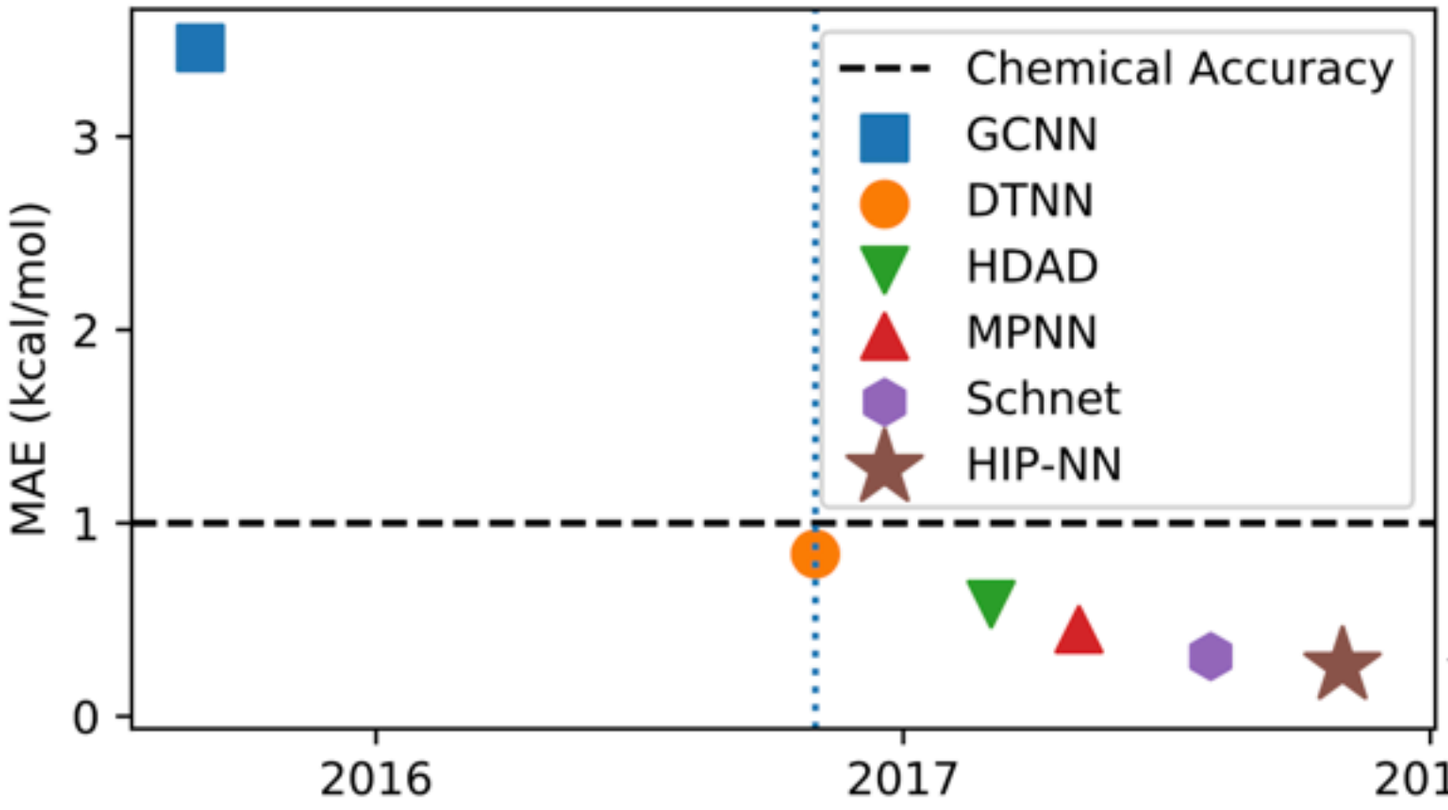
toluene



isoleucine



1 kcal/mol ≈ 1/23 eV

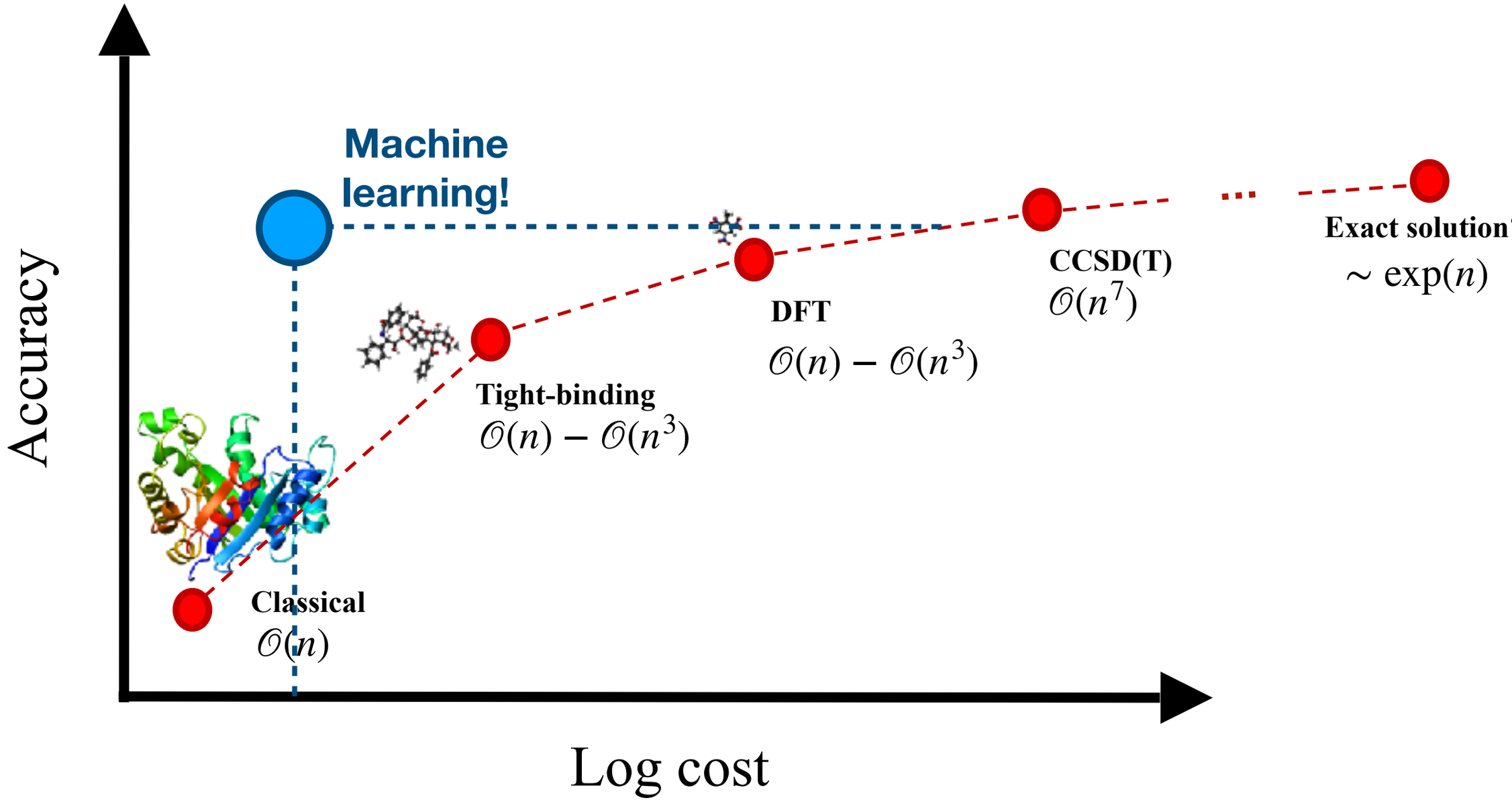


Typical DFT Accuracy

Typical CCSD Accuracy

Lubbers et al., https://arxiv.org/abs/1710.00017

Levels of quantum chemistry



Physics informed machine learning

Symmetries:

Translation Time reversal

Rotation Permutation

Physical principles:

Locality + Coulomb interactions

Hierarchicality (e.g. many-body expansion)

Do *not* constrain to partially correct physics?

Transplant techniques

Deep learning philosophy:

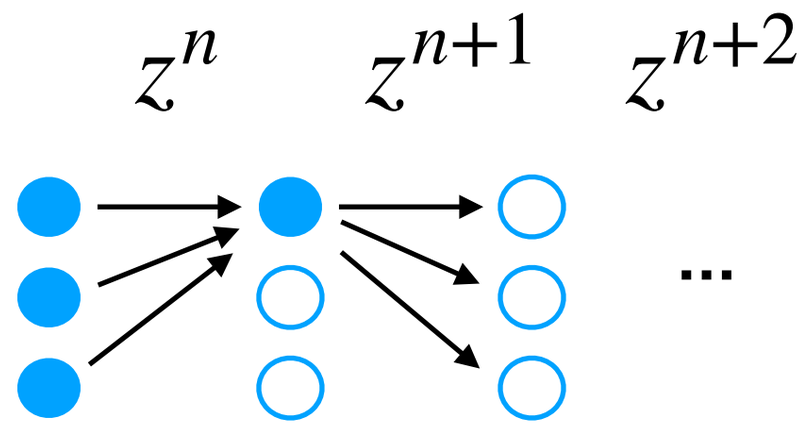
- *don't* hand design features
- use *large* training dataset

VGG: ~10^8 parameters trained on 10^6 images

AlphaGo / AlphaZero

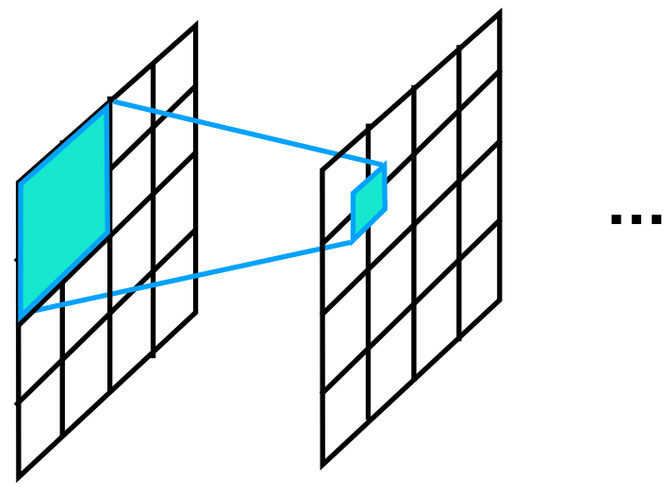
Neural Net

$$z_a^{n+1} = f_{\textbf{nonlin}}(\sum_b W_{ab} z_b^n + b_a)$$



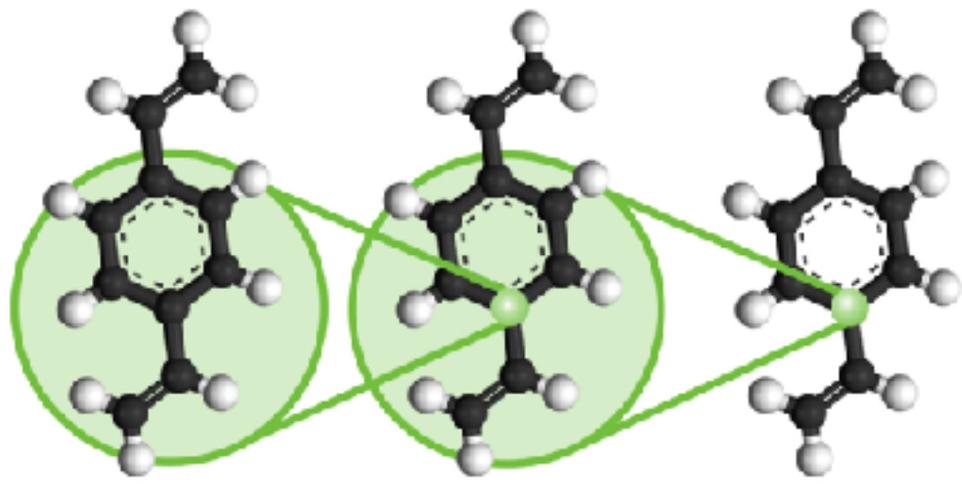
Conv Net

$$z^{n+1}(x) = f_{\textbf{nonlin}}[(W * z^n)(x) + b(x)]$$



Molecular generalization

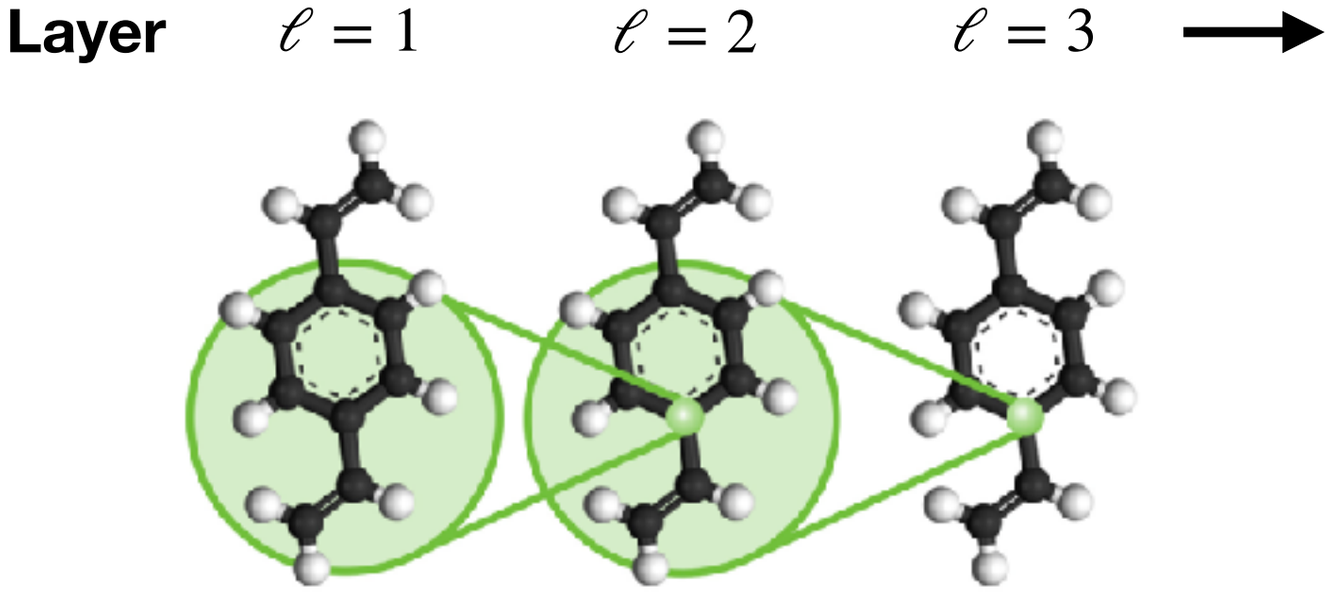
$$z^{n+1}(x) = f_{\textbf{nonlin}}[\ldots z^n \ldots]$$



See also: DTNN, MPNN, SchNet, …

https://ai.googleblog.com/2017/04/predicting-properties-of-molecules-with.html

HIP-NN



Local messages

Pairwise distances
(Translation/rotation invariance)

$$\tilde{z}^{\ell+1}_{i,a} \underset{\text{inter.}}{=} f\left(\sum_{j,b} v^{\ell}_{ab}(r_{ij}) z^{\ell}_{j,b} + \sum_{b} W^{\ell}_{ab} z^{\ell}_{i,b} + B^{\ell}_{a}\right)$$

Descriptors of atom *i*

Learned parameters

Lubbers et al., "Hierarchical modeling of molecular energies using a deep neural network"
https://arxiv.org/abs/1710.00017

Symmetries of ML potential

Translation,
Rotation:

Atomic positions represented entirely with ***pairwise distances***

Time reversal:

Forces are generated as **exact derivative** $f = -\nabla \hat{E}$.

Physical principles

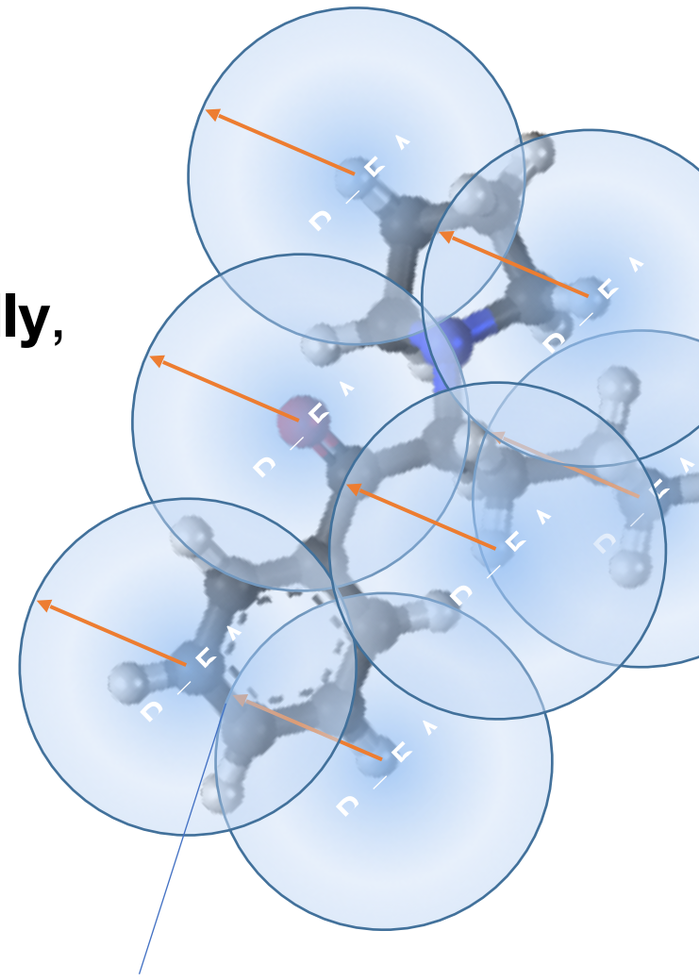
Spatial locality:

Energy decomposed **locally**,

$$\hat{E} = \sum_i \hat{E}_i$$

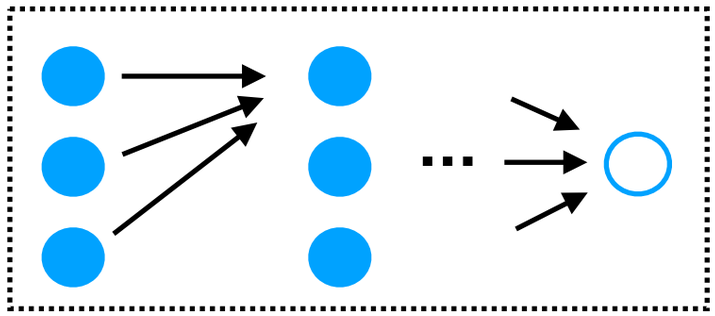
Hierarchicality:

Energy decomposed hierarchically $\hat{E}_i = \sum_\ell \hat{E}_i^\ell$.

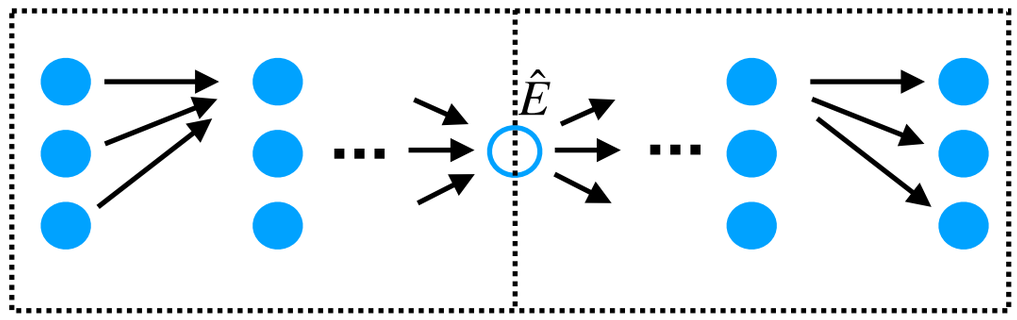


Training to energies *and* forces

Energy $\hat{E}[\mathbf{r}, w]$

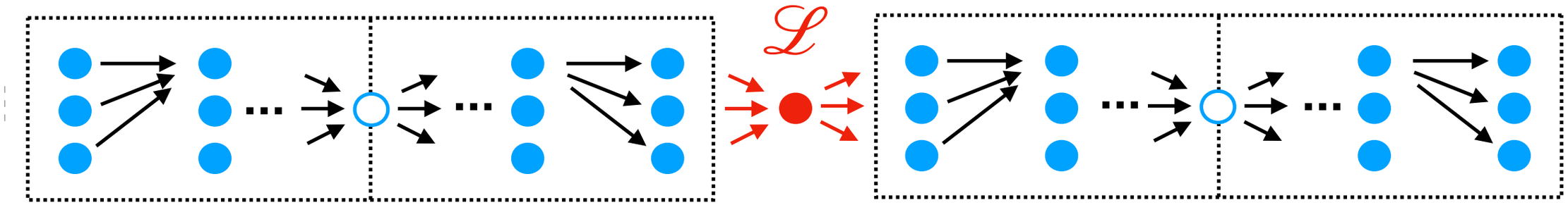


Backprop 1, forces $\nabla_{\mathbf{r}}\hat{E}$



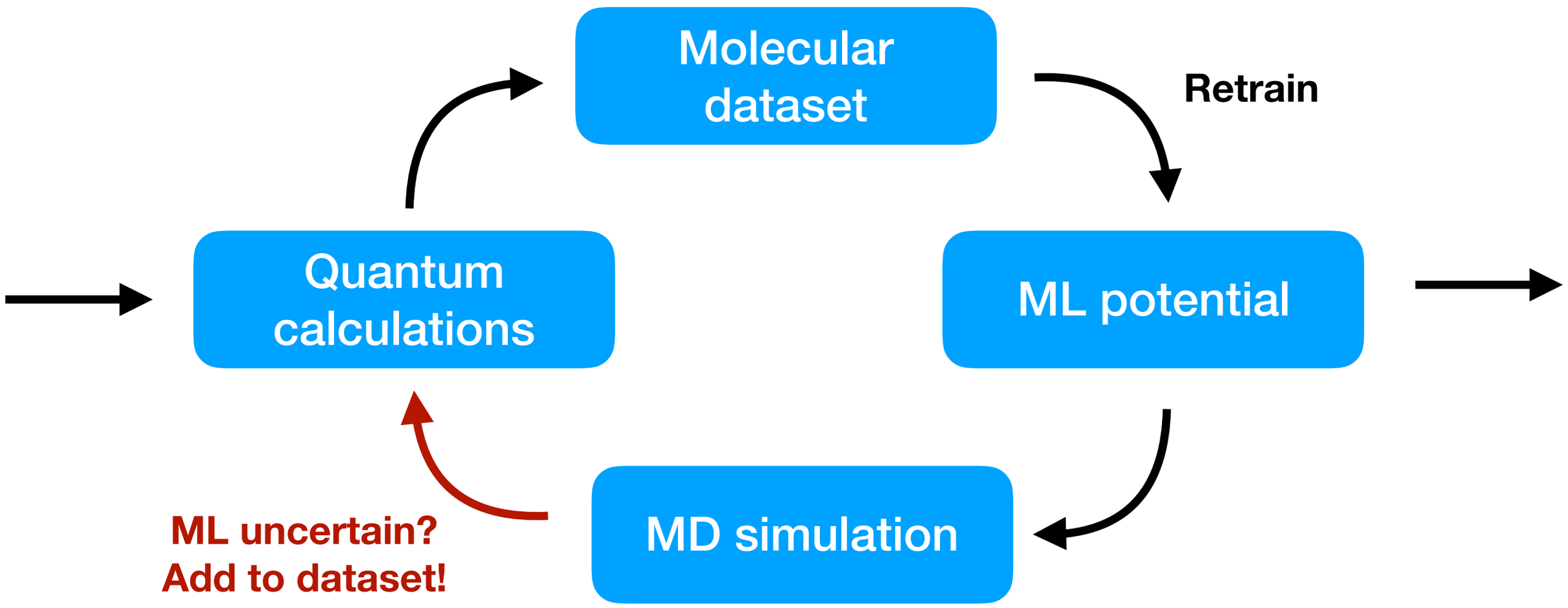
Loss $\mathscr{L} = \sum_{\text{Data}} \left[c_1(E - \hat{E})^2 + c_2(\nabla_{\mathbf{r}}E - \nabla_{\mathbf{r}}\hat{E})^2 \right]$

Backprop 2, weight updates: $\nabla_w \mathscr{L}$



Active learning

Computer should help us design good datasets!

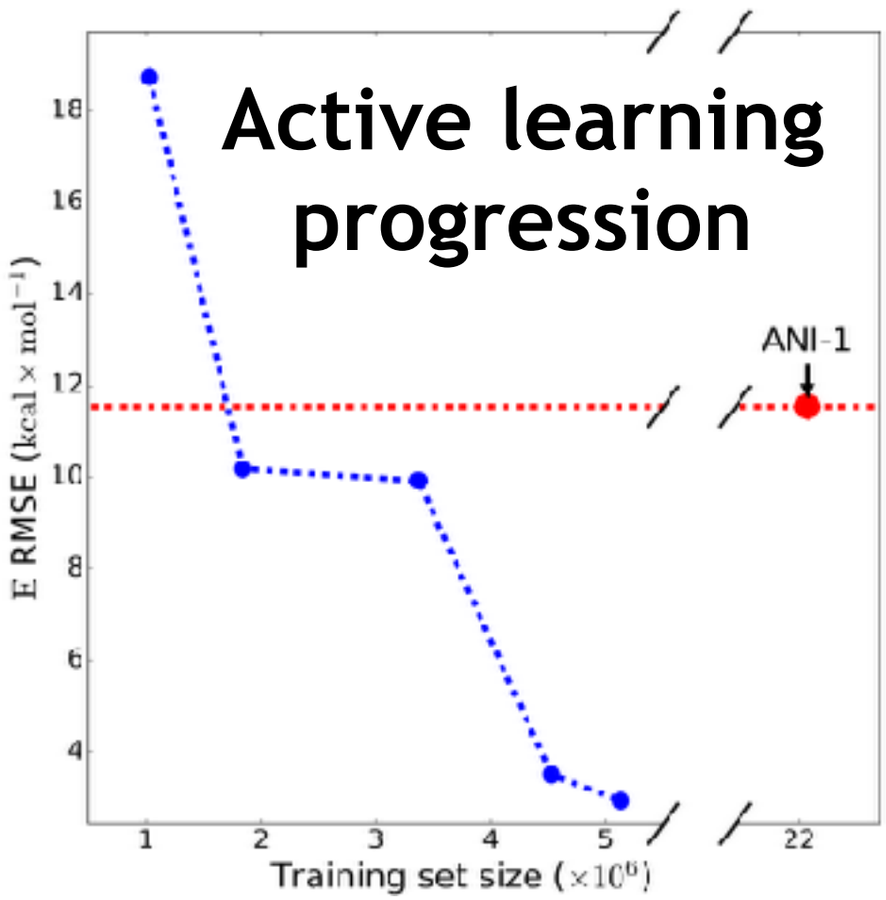


Smith et al., “Less is more: sampling chemical space with active learning”
https://arxiv.org/abs/1801.09319

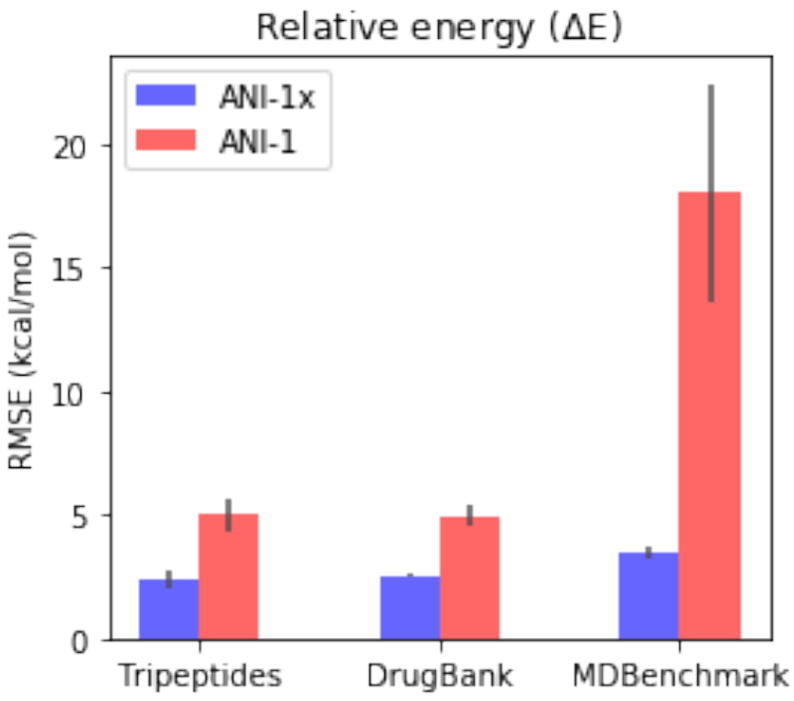
Active learning vs random sampling

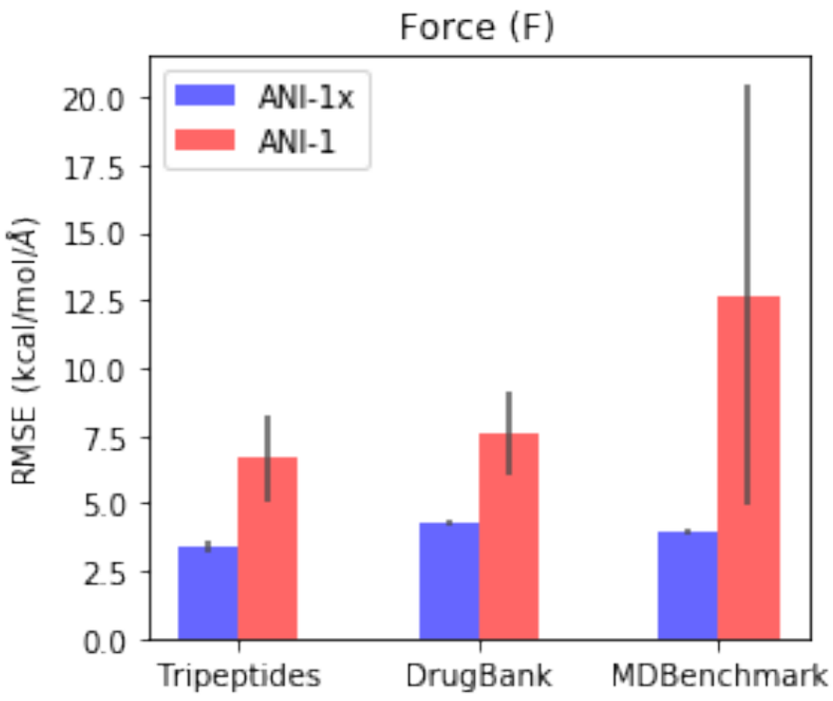
Dataset sizes

ANI-1	ANI-1x
22M	5M



Errors



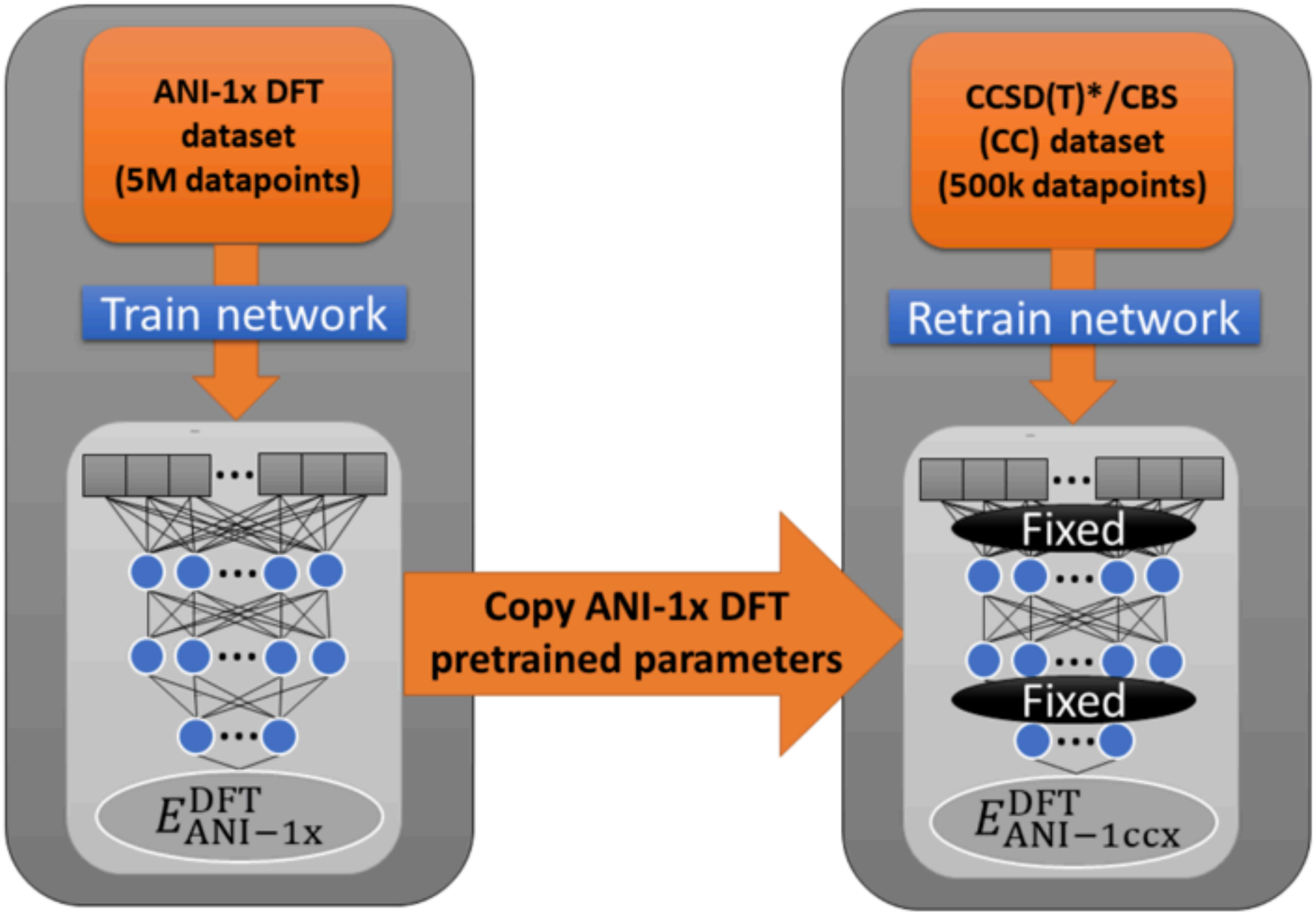


JS Smith, et al.; *The Journal of Chemical Physics*, (**2018**), 148 (24), 241733

Transfer learning

Combine lots of DFT data with some high accuracy CCSD data

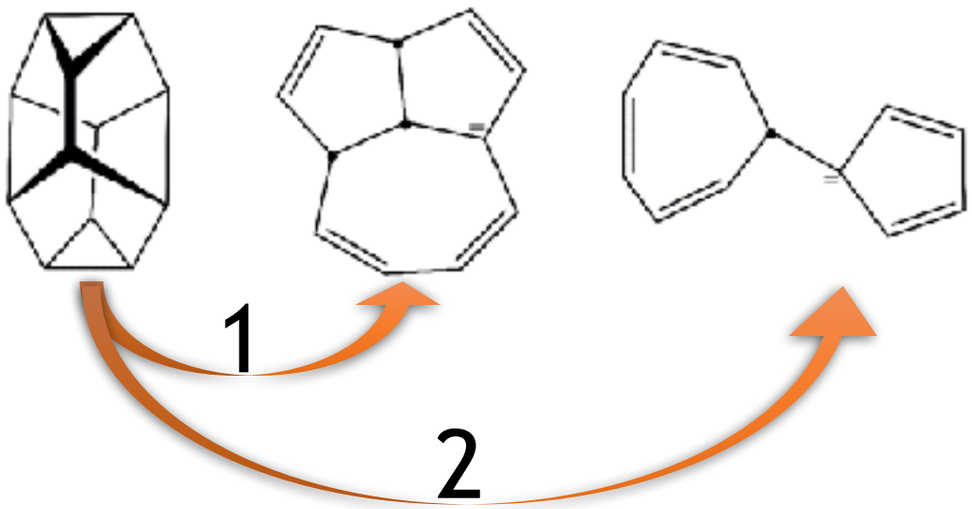
- Subsample 10% of ANI-1x training data (0.5M of 5M)
- Recompute CCSD(T)/CBS level
- 340k parameters fixed, re-train 60k
- 10^7 faster than DFT

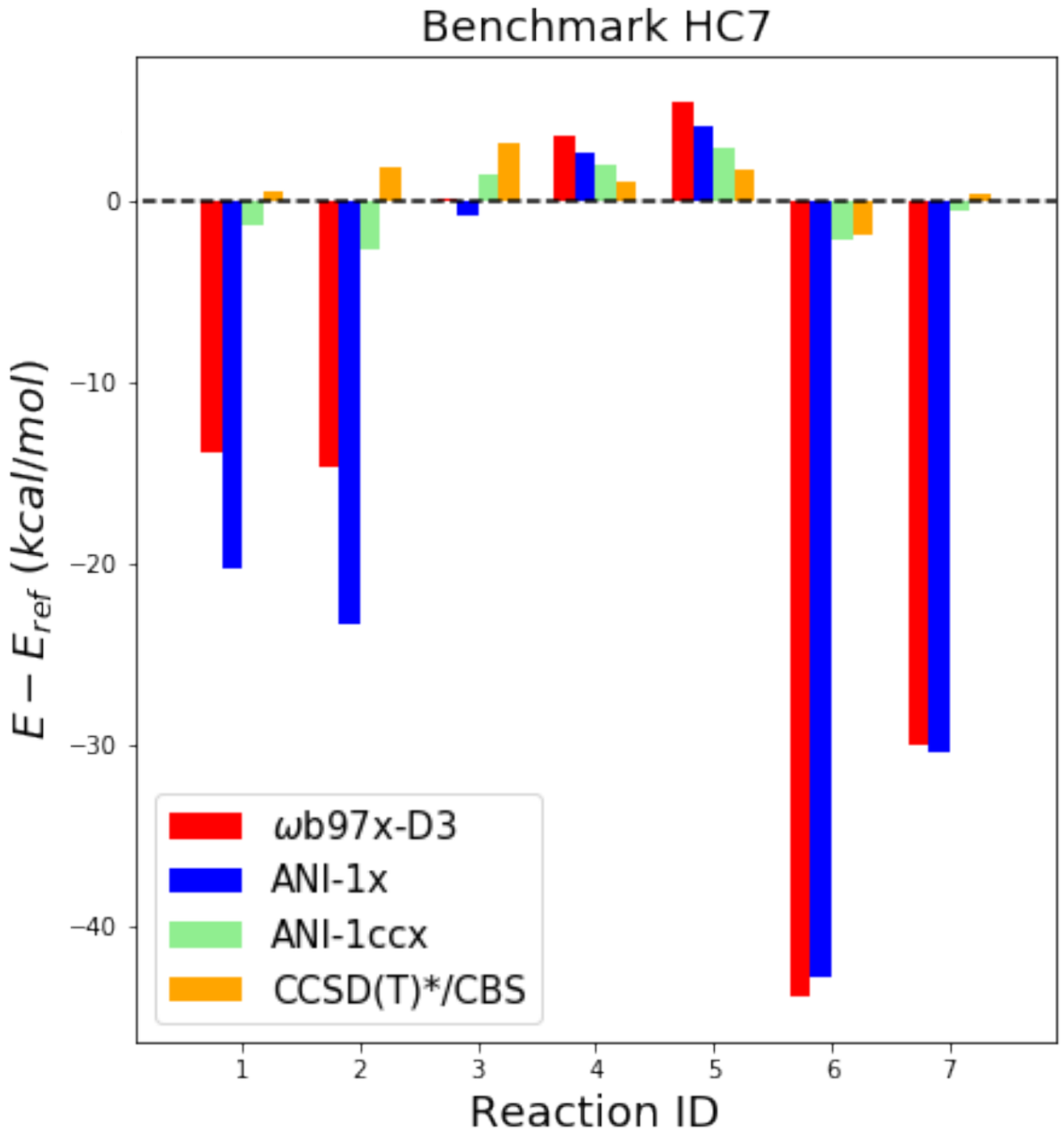


Smith et al., "Outsmarting Quantum Chemistry Through Transfer Learning" [ChemRxiv:6744440]

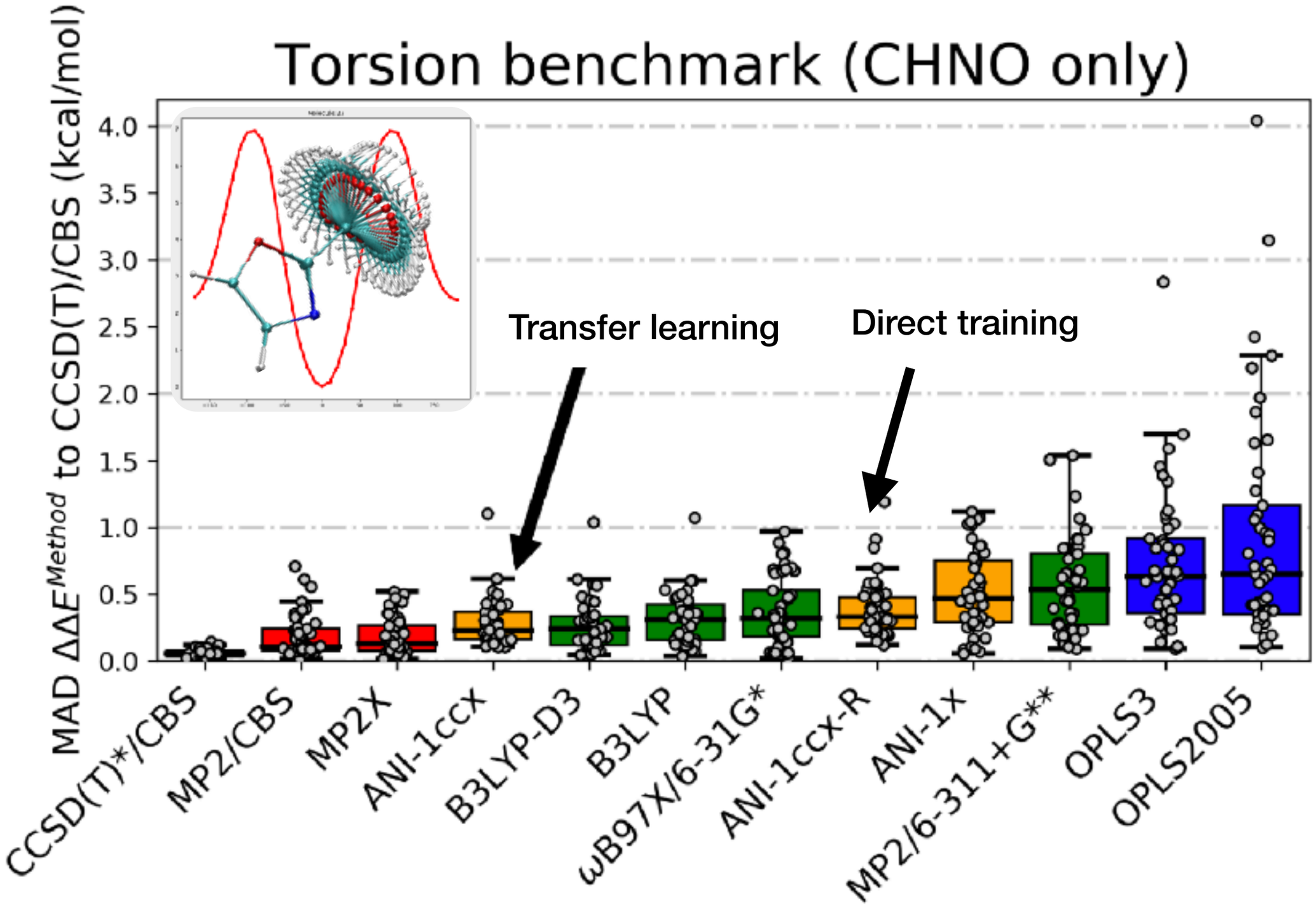
Hydrocarbon reaction energy benchmark

Examples





Reference data: Peverati, R.; Zhao, Y.; Truhlar, D. G., *J. Phys. Chem. Lett.* **2011**, *2* (16), 1991–1997.

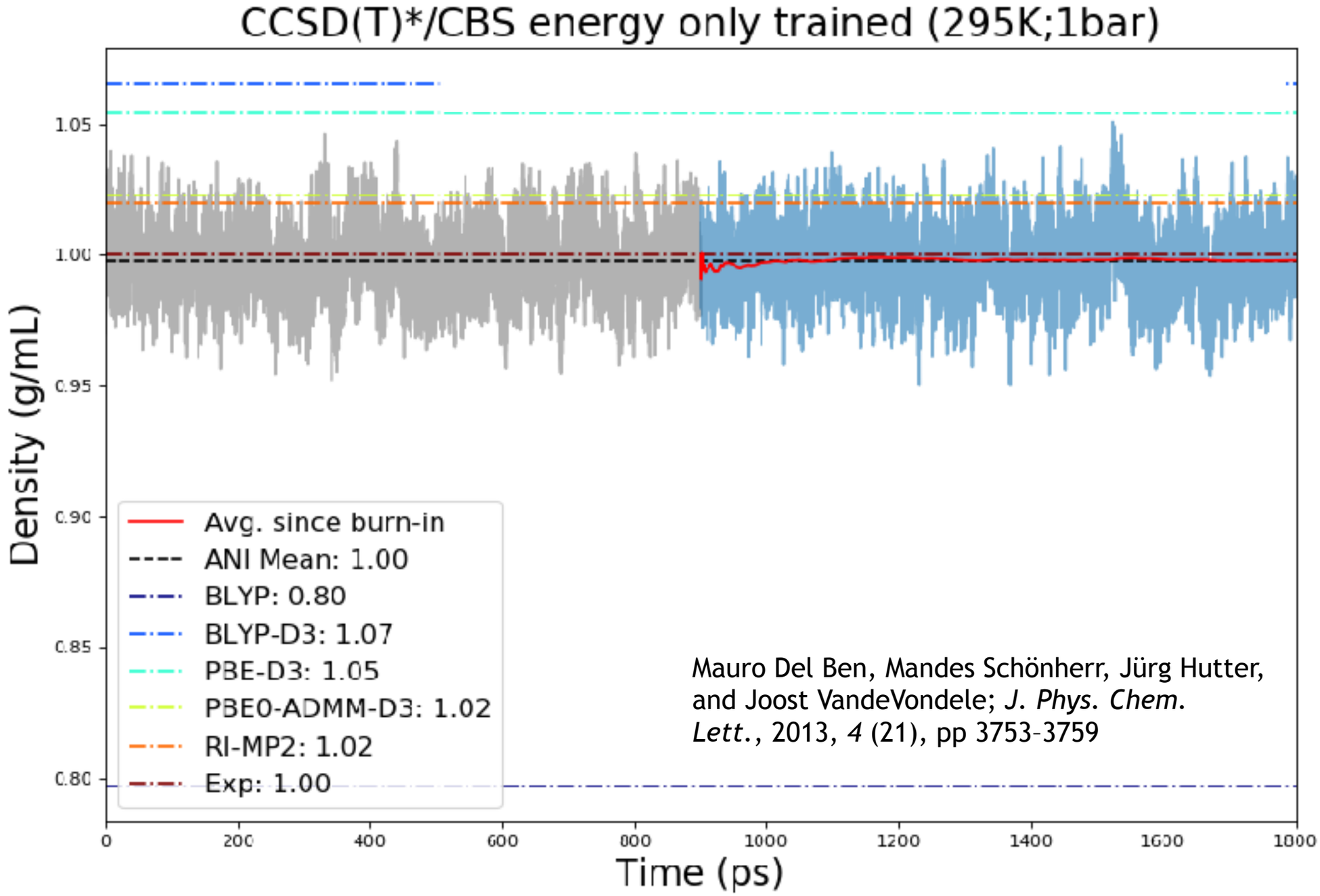


Sellers, B. D.; James, N. C.; Gobbi, A. A Comparison of Quantum and Molecular Mechanical Methods to Estimate Strain Energy in Druglike Fragments. *J. Chem. Inf. Model.* **2017**, *57* (6), 1265–1275.

Smith et al., “Outsmarting Quantum Chemistry Through Transfer Learning” [ChemRxiv:6744440]

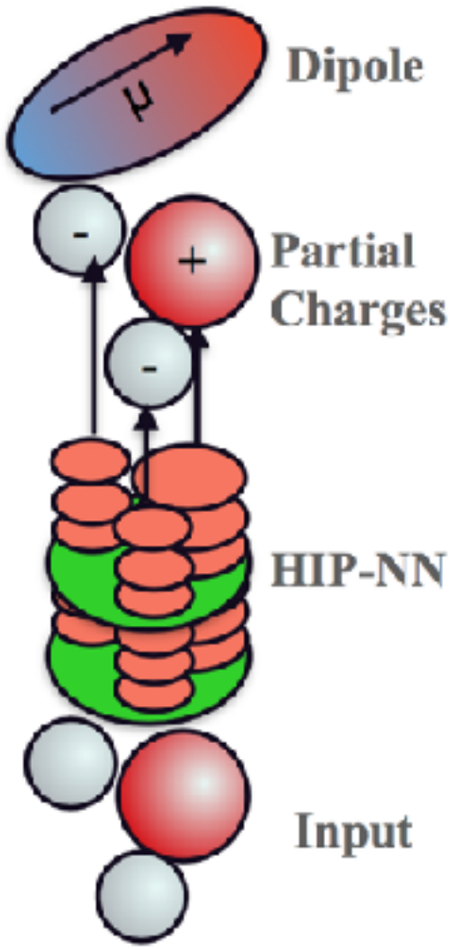
Density of bulk water

NPT conditions: 295K; 1Bar
267 waters

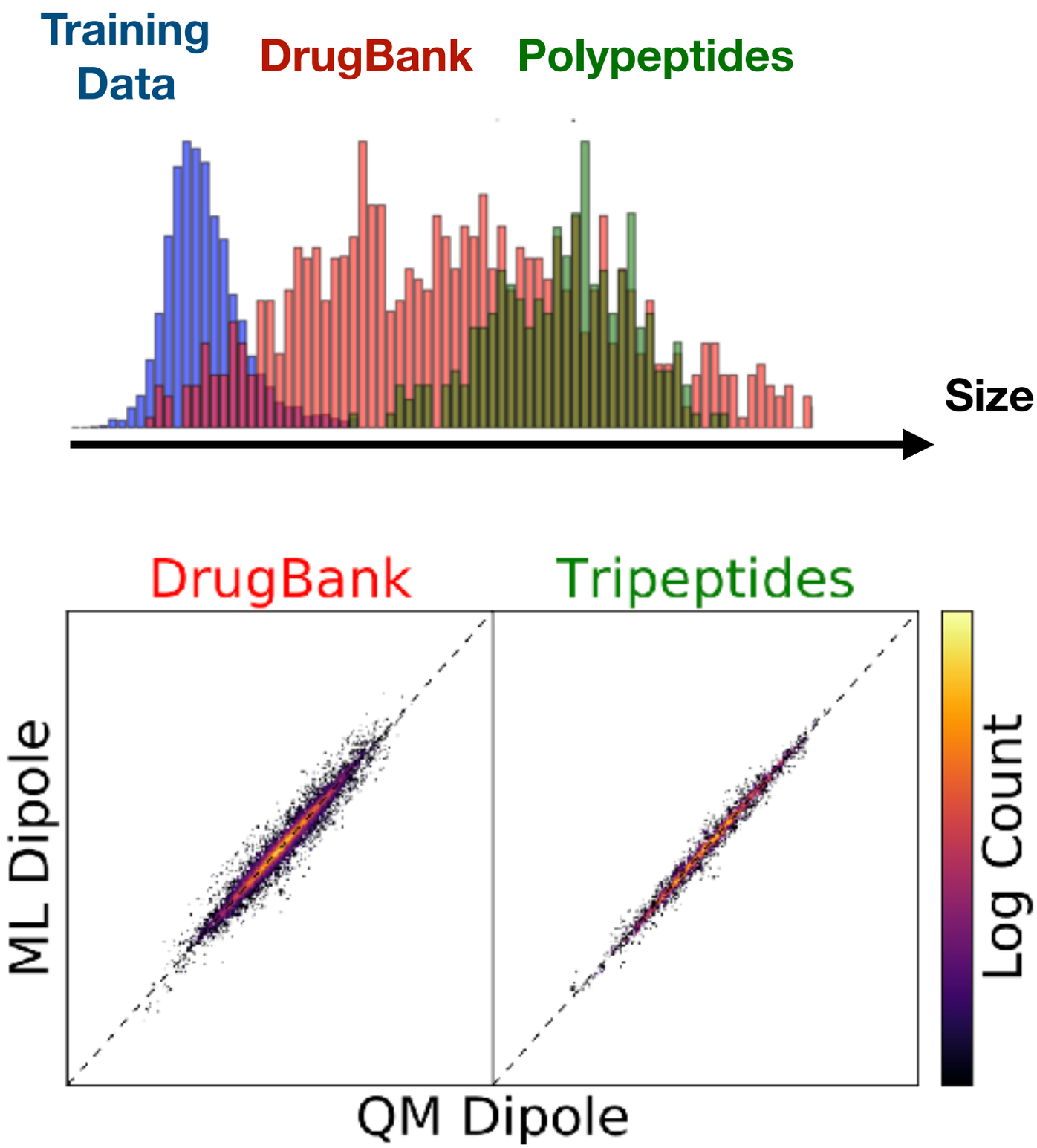


J. S. Smith, in preparation

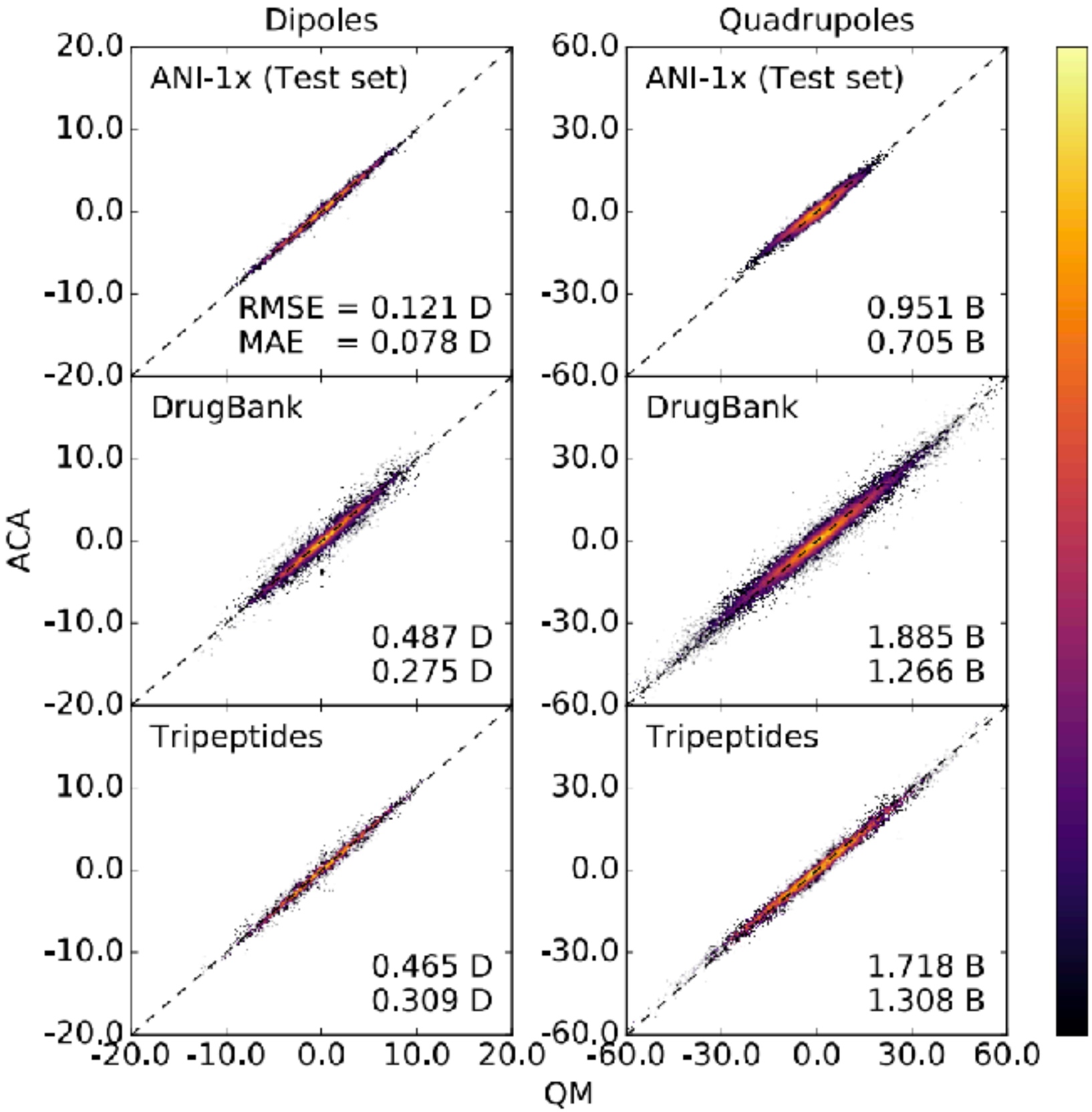
Inferring local charges to produce dipoles



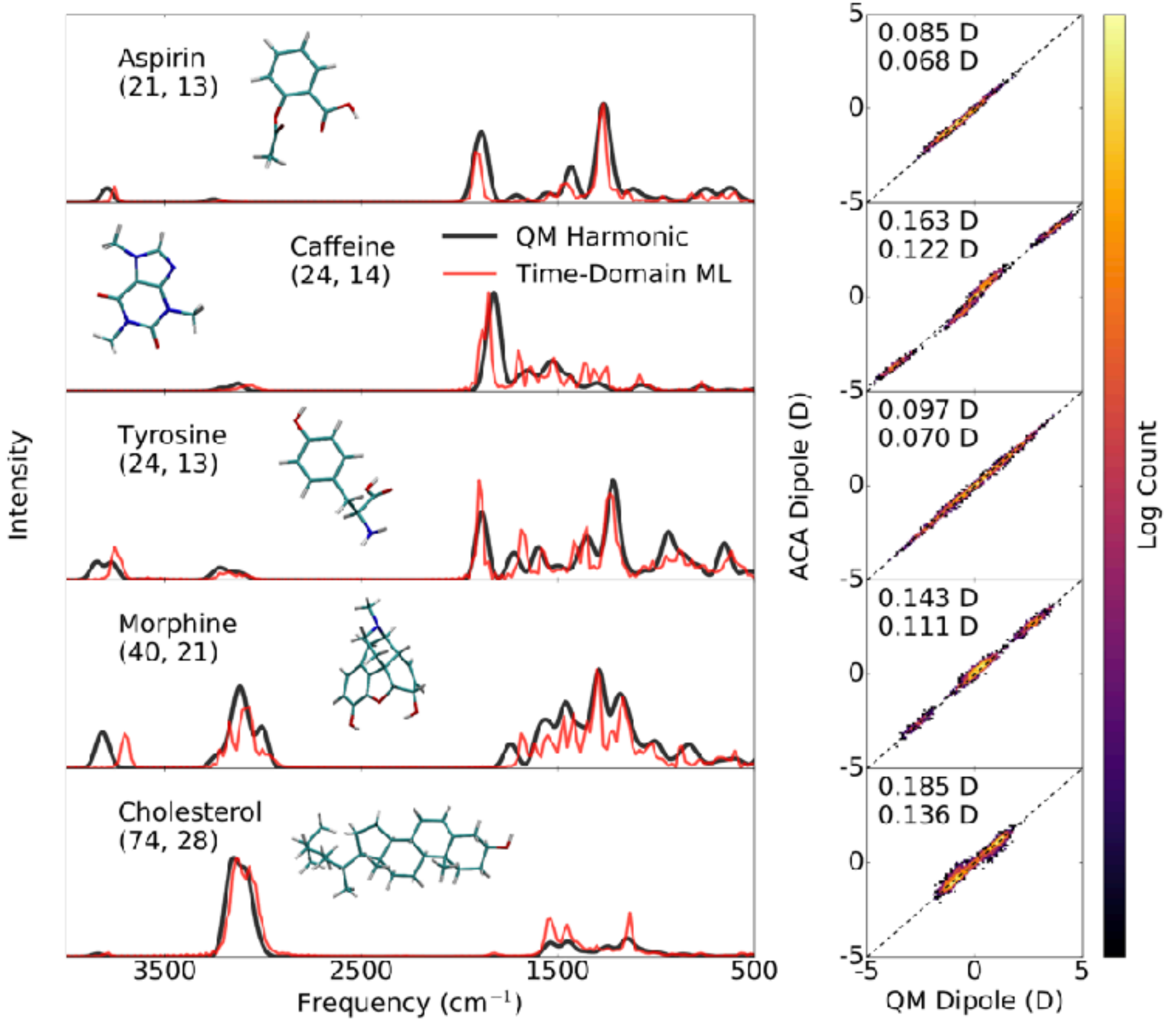
Sifain et al., "*Discovering a Transferable Charge Assignment Model Using Machine Learning*" [ChemRxiv:6638981]



Inferred charges generalize to *quadrupoles*



IR Spectra


Aspirin
(21, 13)
Caffeine
(24, 14)
Tyrosine
(24, 13)
Morphine
(40, 21)
Cholesterol
(74, 28)
QM Harmonic
Time-Domain ML
Intensity
3500
2500
1500
500
Frequency (cm^{-1})
0.085 D
0.068 D
0.163 D
0.122 D
0.097 D
0.070 D
0.143 D
0.111 D
0.185 D
0.136 D
ACA Dipole (D)
QM Dipole (D)
Log Count

Conclusions

Machine learning to *emulate quantum chemistry* works great!

Future directions?

Accounting for electron *dynamics*?

Joint-training to more information from wave-function?

Use ML to predict effective quantum Hamiltonians?

Better transfer learning? Can we incorporate data from experiment (e.g. phase diagrams)?

Kipton Barros

kbarros@gmail.com

cnls.lanl.gov/~kbarros

Justin Smith

Nicholas Lubbers

Lubbers et al., "*Hierarchical modeling of molecular energies using a deep neural network*" [arXiv:1710.00017]

Smith et al., "*Less is more: sampling chemical space with active learning*" [arXiv:1801.09319]

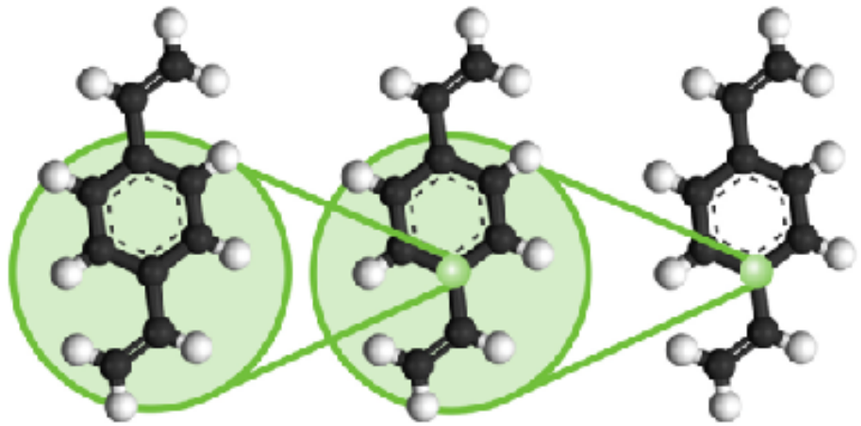
Smith et al., "*Outsmarting Quantum Chemistry Through Transfer Learning*" [ChemRxiv:6744440]

Sifain et al., "*Discovering a Transferable Charge Assignment Model Using Machine Learning*" [ChemRxiv:6638981]

Entire HIP-NN

$$E \approx \hat{E} = \sum_{i=1}^{N_{\text{atom}}} \hat{E}_i. \qquad \hat{E}_i = \sum_{\ell} w_a^{\ell} z_{i,a}^{\ell}$$

Locality/ Hierarchicality



$$f(x) = \log(1 + e^x).$$

Activation fn.

$$z_{i,a}^{\ell+1} \underset{\text{inter.}}{=} f\left(\sum_{j,b} v_{ab}^{\ell}(r_{ij}) z_{j,b}^{\ell} + \sum_{b} W_{ab}^{\ell} z_{i,b}^{\ell} + B_a^{\ell}\right)$$

Interactions

$$z_{i,a}^{\ell+1} = \sum_{b} \left(\tilde{W}_{ab}^{\ell} \tilde{z}_b^{\ell+1} + \tilde{M}_{ab}^{\ell} z_{i,b}^{\ell}\right) + \tilde{B}_a^{\ell},$$

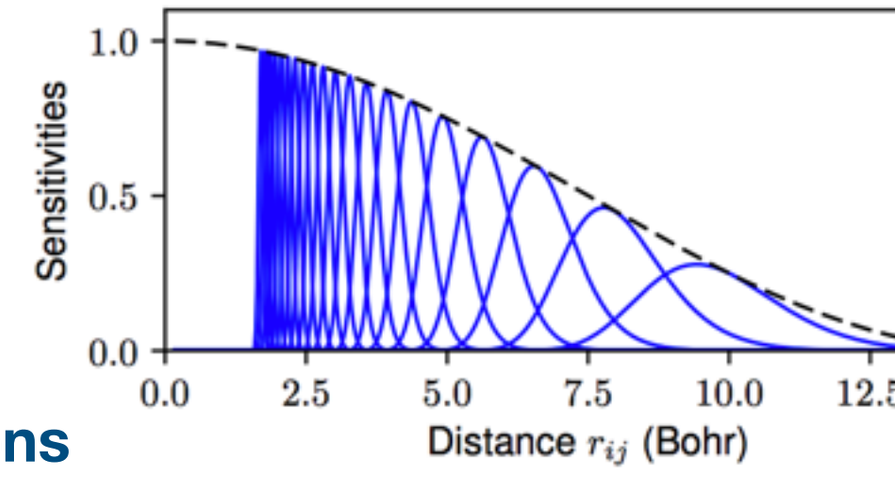
Res-net

$$v_{ab}^{\ell}(r_{ij}) = \sum_{\nu} V_{\nu,ab}^{\ell} s_{\nu}^{\ell}(r_{ij}),$$

Basis expansion

$$s_{\nu}^{\ell}(r) = \exp\left[-\frac{\left(r^{-1} - \mu_{\nu,\ell}^{-1}\right)^2}{2\sigma_{\nu,\ell}^{-2}}\right] \varphi_{\text{cut}}(r).$$

Sensitivity functions



$$\mathscr{L} = \sum_{\textbf{Data}} \left[c_1(E - \hat{E})^2 + c_2(\nabla E - \nabla \hat{E})^2\right]$$

Loss fn.